

The New Science of Causation

It seems like each new day brings another one of those headlines: regular sleep "linked to" life expectancy, playing video games "associated with" surgical prowess, bullies "at risk" of becoming criminals, and "does breastfeeding reduce a baby's blood pressure?" (the old rhetorical question gambit). Sometimes the articles are clear: the research has only found a correlation between two variables — breastfeeding and low blood pressure were found together. But more often, they imply that causation is at work — that breastfeeding *causes* lower blood pressure.

You've surely heard that old statistics adage: correlation does not imply causation. Just because breastfeeding and low blood pressure are found together doesn't mean the first caused the second. Perhaps the second caused the first (moms might prefer to keep breastfeeding calmer babies) or some other thing caused both of them (maybe moms who don't work both tend to breastfeed and stress their kids less). You can't tell from correlation alone.

Indeed, the philosopher David Hume argued that we could *never* know whether causation was at work. "Solidity, extension, motion; these qualities are all complete in themselves, and never point out any other event which may result from them," he wrote. But not causation: "One event follows another; but we never can observe any tie between them. They seem conjoined, but never connected."

And, as philosophers since Plato and Sextus Empiricus have argued, such evidence can deceive us. Imagine finding a button and, each time you press it, a beep is heard. Normally, we'd assume that the button always causes the beep. But we'd be wrong — one day the power goes out and the button does nothing.

Which is why, centuries later, Karl Pearson, the founder of mathematical statistics, banned the notion of causality from the discipline, calling it "a fetish amidst the inscrutable arcana of modern science" and insisting that just by understanding simple correlation one "grasped the essence of the conception of association between cause and effect."

His followers have kept it banished ever since. "Considerations of causality should be treated as they have always been in statistics: preferably not at all," wrote a former president of the Biometric Society. "It would be very healthy if more researchers abandon thinking of and using terms such as cause and effect," insisted another prominent social scientist.

And there the matter has stayed. Causality is a concept as meaningless as "the soul" and just as inappropriate for modern mathematical science. And yet, somehow, this doesn't seem quite right. If causation is nothing but a meaningless word that laypeople have layered over correlation, then why the ceaseless insistence that "correlation does not imply causation"? Why are our thoughts filled with causal comments (he made me do it!) and never correlational ones?

The result is exceptionally strange. Statistics has no mathematical way to express the notion "mud does not cause rain". It can say mud is correlated with rain (i.e. that there's a high probability of seeing mud if you see rain), no problem, but expressing the simple causal concept — the kind of thing any five-year-old would know — is impossible.

Statisticians may have never had to confront this problem but, luckily for us, Artificial Intelligence researchers have. It turns out if you're making a robot, having a notion of causality is essential — not just because it's the only way to understand the humans, but because it's the only way to get anything done! How are you supposed to turn the lights on if you don't know that it's the light-switch and not the clicking noise that causes it?

The result is that in recent years several teams of AI researchers have turned their focus from building robots to building mathematical tools for dealing with causality. At the forefront is Judea Pearl (author of the book *Causality*, Cambridge University Press) and his group at UCLA and Clark Glymour (author of *The Mind's Arrows*, MIT Press), Peter Spirtes, and their colleagues at Carnegie Mellon. The result is a quiet revolution in the field of statistics — one most practicing statisticians are still unaware of.

They started by dismissing Plato's skepticism about the problem. Granted, they say, we may never know for sure whether the button *always* causes the beep, but that's too stringent a demand. Science never knows anything for sure — the best we can hope for is extracting the most knowledge from the evidence we have. Or, as William James put it, "To know is one thing, and to know for certain that we know is another."

Next, they created a new mathematical function to formalize our notion of causality: $do(\dots)$. do expresses the notion of intervening and actually trying something. Thus, to mathematically express the notion that mud does not cause rain, we can say $P(\text{rain} | do(\text{mud}=\text{true})) = P(\text{rain})$ — in other words, the chance of rain given that you made it muddy is the same as the chance of rain in general.

But causes rarely comes in pairs like these — more often it comes in complicated chains: clouds cause rain which causes both mud and wet clothing and the latter causes people to find a change of clothes. And so the researchers express these as networks, usually called causal Bayes nets or graphical causal models, which show each thing (clouds, rain, mud) as a node and the causal relationships as arrows between them:

And all this was just the warm-up act. Their real breakthrough was this: just as kids can discover causes by observation, computers can discern causes from data. Now obviously the easiest way to do this is just to measure what happens when you $do(X=x)$ directly — this, for example, is what randomized controlled trials do. Kids do it by dropping a fork on the floor and seeing if this causes Mom to pick it up; scientists do it by randomly giving some people a real drug and others just a placebo. The result is that we can be sure of the cause — after all, it was *we* who dropped the fork and gave out the drug; nothing else could be sneaking in and causing it.

But in most cases we don't have this luxury. We'd like to know whether a new tax policy will cause the economy to tank *before* we enact it; we'd like to know whether smoking causes cancer without forcing kids to smoke; and even in randomized controlled trials, we can give half the patients the real drug, but we can't make them take it. If the drug being tested makes someone so horribly sick that they stop taking it and then get better, drug trials still count that as a victory for the drug!

Obviously we can't always know such things just from observing, but in a surprising number of cases we can. And the researchers have developed a mathematical method — called the do -calculus — for determining just when you can. Feed it a Bayes network of variables, their relationships, and their values, and it will return back what it knows and with what certainty.

Thus, in an example Pearl frequently uses, tobacco companies used to argue that the correlation between smoking and cancer was simply because there were certain genes that made people both more likely to smoke *and* more likely to get cancer. It didn't matter if they quit smoking — their genes would lead cancer to get them anyway. Pearl shows that if we assume only smoking causes tar deposits on the lungs and the tar deposits are the only way smoking causes cancer, we can simply measure the tar deposits and calculate whether the tobacco companies are right.

Or, in another example in his book *Causality*, he analyzes data from a study on a cholesterol-reducing drug. Since whether people got the placebo or not is unassociated with any other variables (because it was randomly assigned) if we merely assume that receiving the real drug has *some* influence on whether people take it, we can calculate the effectiveness of the drug even with imperfect compliance. Indeed, we can even estimate how effective the drug would have been for people who were assigned it but didn't take it!

And that's not all — Peter Spirtes and Clark Glymour have developed an algorithm (known as PC, for Peter-Clark) that, given just the data, will do its best to calculate the causal network behind it. You can download the software implementing it, called [TETRAD IV](#), for free from their department's website — it even has a nice graphical interface for drawing and displaying the networks.

As an experiment, I fed it some data from the IRS about 2005 income tax returns. It informed me that the percentage people donate to charity is correlated with the number of dependents they have, which in turn correlates with how much people receive from EITC. That amount, along with average income, causes how many people are on EITC. Average income is correlated with the tax burden which is correlated with inequality. All interesting and reasonable — and the result of just a few minutes' work.

The applications for such tools are endless. As Pearl points out, they have the possibility to radically improve how statistics are used in medicine, epidemiology, economics, sociology, and law. And, as Glymour observes, it lets us better understand results in neuroscience and psychology. Take *The Bell Curve*, the 1992 bestseller that argued blacks had lower IQs, causing poorer performance in school and thus lower-paying jobs and more crime. Glymour shows, by applying the *do*-calculus, these results only hold if you assume that there are no other interactions between the variables (e.g. that parental attitude toward learning doesn't affect both IQ and performance in school). But the PC algorithm and TETRAD IV can demonstrate otherwise.

Such results may be a revolution in social science, but compared to building human-like robots, they're child's play. That's certainly the impression one gets from Pearl. Discussing his work at a conference of Artificial Intelligence researchers, he said:

One of the reasons I find these areas to be fertile grounds to try out new ideas is that, unlike AI, tangible rewards can be reaped from solving relatively small problems. Problems involving barely 4 to 5 variables, which we in AI regard as toy-problems, carry tremendous payoffs in public health and social science.

Billions of dollars are invested each year on various public-health studies: Is chocolate ice-cream good for you or bad for you? Would red wine increase or decrease your heart rate? etc. etc.

The same applies to the social sciences. Would increasing police budget decrease or increase crime rates? Is the Colorado school incident due to TV violence or failure of public education? The Inter-university Consortium for Political and Social Research has distributed about 800 gigabytes worth of such studies in 1993 alone.

Unfortunately the causal-analytical methodology currently available to researchers in these fields is rather primitive, and every innovation can make a tremendous difference. [...] This has been changing recently as new techniques are beginning to emerge from AI laboratories. I predict that a quiet revolution will take place in the next decade in the way causality is handled in statistics, epidemiology, social science, economics, and business. While news of this revolution will never make it to DARPA's newsletter, and even NSF is not equipped to appreciate or support it, it will nevertheless have enormous intellectual and technological impact on our society.

For science's sake, I hope he's right.

You should follow me on twitter [here](#).

September 21, 2009